(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 126 439 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
22.08.2001 Bulletin 2001/34

(51) Int Cl.7: **G10L 21/02**

(21) Application number: **00307402.8**

(22) Date of filing: **29.08.2000**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **14.02.2000 US 503990**

(71) Applicant: **LUCENT TECHNOLOGIES INC.**
Murray Hill, New Jersey 07974-0636 (US)

(72) Inventors:
- **Sherif, Mahmoud R.**
  Bayonne, New Jersey 07002 (US)
- **Tarraf, Ahmed A.**
  Bayonne, New Jersey 07002 (US)

(74) Representative:
Watts, Christopher Malcolm Kelway, Dr. et al
Lucent Technologies (UK) Ltd,
5 Mornington Road
Woodford Green Essex IG8 OTU (GB)

(54) Mobile to mobile digital wireless connection having enhanced voice quality

(57) This invention is directed toward improving the subjective voice quality of a mobile to mobile phone call having tandem vocoder processing by modifying the spectrum of the voice signal before it is processed by the second vocoder to compensate for digital distortion which is generated by the second vocoder. An adaptive filter can be used to modify the spectrum of the voice signal. With this invention, the voice quality of a call from a first mobile phone to a second mobile phone has a quality that is substantially similar to the voice quality of a call from a mobile phone to a desk phone.

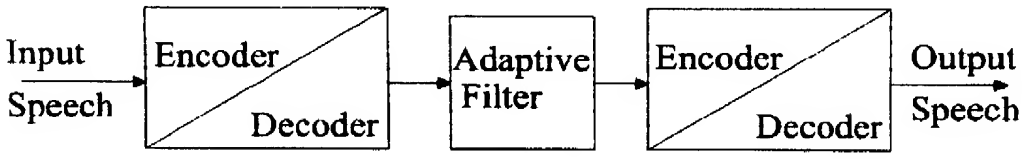


FIG. 4

EP 1 126 439 A2

Description

BACKGROUND OF THE INVENTION

1. Field of the Invention

[0001] The present invention relates generally to wireless communication, and more particularly to mobile to mobile wireless communications connections having improved voice quality.

2. Description of the Related art

[0002] Wireless telephone communication systems utilize encoding to limit the amount of bandwidth required to transmit voice information over a communication channel. Instead of sending the received voice over the channel, a voice coder (vocoder) is used to first analyze the received voice signal and then generate a set of parameters which are used to drive a model that synthesizes a replica of the voice. It is these parameters that are transmitted at very low rates over the telephone network. Typically, in one type of vocoder, the voice is analyzed on a short time basis and a set of model parameters which may represent pitch, line spectrum frequencies (the frequency response of the vocal tract at any given time) and voice gain are generated.

[0003] In a digital wireless network, a vocoder is used at the transmitter to divide the voice signal of the person talking into a stream of fixed length frames of typically 20 msec. The vocoder then models the voice using a specific mathematical model and for each 20 msec frame, the vocoder generates a set of parameters for that mathematical model that best describes the voice for that 20 msec interval. Thus, the vocoder encodes the voice of the speaker into 20 msec frames. Each one of the frames carries a set parameter that represents the speaker's voice for the corresponding 20 msec interval. A modem at the transmitter transmits these frames over the air using an air interface technique such as CDMA, TDMA, or GSM. The modem function at the receiver passes the frames to a decoder at the receiver. The decoder decodes the 20 msec frames to produce the voice of the subscriber. A wireless subscriber uses a wireless terminal to access the wireless network.

[0004] The audio path from a mobile station to a desktop phone consists of a voice encoder at the mobile station, a wireless transmitter at the mobile station, a wireless receiver at the base station, and a decoder at the base station. There is one voice encoder and one decoder involved in the processing of the audio signal. This process is referred to as being a single vocoding operation because it involves a single encoding/decoding process.

[0005] The audio path from a first mobile station to a second mobile station consists of a voice encoder and a wireless transmitter at the first mobile station, and a wireless receiver and a voice decoder at the base station (or at the switch). The voice signal of the subscriber generated by the decoder is transmitted over the carrier's land lines to a second voice encoder at the second base station (or at the switch), and then to a wireless transmitter. A wireless receiver at the second mobile station receives the encoded voice signal and processes it in a second decoder to generate a voice signal. In this case, from the voice processing perspective, there is one voice encoder at the first mobile station, one decoder at the base station, another voice encoder at the second base station and a second decoder at the receiving mobile station. Thus, when a call is made from one mobile station to another mobile station, the voice signal is processed by two voice encoders and two decoders. This process is referred to as tandem vocoding processing because the voice signal is encoded twice and decoded twice.

[0006] The processing of the voice signal (the input speech) by the vocoding process (the process of encoding and decoding the voice signal) introduces a certain amount of digital distortion into the voice generated at the receiver relative to the actual voice uttered by the speaker. While, the quality of the voice generated by the decoder in single vocoding processing is very acceptable, it is somewhat below the voice quality of the input speech. Tandem vocoding compounds the digital distortion introduced into the voice signal and, therefore, it has been observed that the subjective voice quality of a mobile to mobile call is noticeably below that of a mobile to desk phone call. Thus, there exists a need to provide a method and apparatus for enhancing the voice quality of a mobile to mobile call.

Summary of the Invention

[0007] This invention is directed toward improving the subjective voice quality of a mobile to mobile phone call having tandem vocoder processing by modifying the spectrum of the voice signal before it is processed by the second vocoder to compensate for digital distortion which is generated by the second vocoder. An adaptive filter can be used to modify the spectrum of the voice signal. With this invention, the voice quality of a call from a first mobile phone to a second mobile phone has a quality that is substantially similar to the voice quality of a call from a mobile phone to a desk phone.

BNSDOCID: <EP_____1126439A2_I_>

Brief Description of the Drawing

[0008]

Fig. 1 is a block diagram of a mobile to desk phone call having single vocoding processing;
Fig. 2 is a block diagram of a mobile to mobile call having double vocoding processing;
Fig. 3 illustrates the spectrum of the output speech produced during single vocoding processing (graph A) and double vocoding processing (graph B) for the same phonetically balanced input voice signal;
Fig. 4 is a block diagram of one embodiment of structure in accordance with the principles of the invention;
Fig. 5 is a block diagram of another embodiment of structure in accordance with the principles of the invention;
Fig. 6 is a block diagram of still another embodiment of structure in accordance with the principles of the invention;
Fig. 7 illustrates the difference in the spectrum of two signals, one that is subjected to tandem vocoding processing and the other that is subjected to single vocoding processing;
Fig. 8 illustrates the frequency response of an adaptive filter which can be located between two vocoders in tandem to compensate for digital distortion generated by the second vocoder;
Fig. 9 illustrates the spectrum of a voice signal processed by tandem vocoding including the adaptive filter of Fig. 7 relative to the spectrum of a voice signal processed by single vocoding; and
Fig. 10 illustrates the plot of the frequency response of an adaptive filter for use with two EVRC vocoders in tandem.

Detailed Description of the Invention

[0009] As noted previously, in a mobile to desk phone conversation, the speech signal goes through a single vocoding process which introduces a small amount of digital distortion into the produced speech. Thus, the quality of the produced speech is somewhat less than that of the input speech. For a call between two mobile stations where the voice signal is subjected to the vocoding process twice, the subjective voice quality of the produced speech experiences increased digital distortion and, therefore, is inferior to speech produced by a voice signal that had experienced only single vocoding processing.

[0010] Referring to Table A below, there is provided the Mean Opinion Score of the quality of a speech signal generated with single vocoding processing and tandem vocoding processing for various vocoders. The Mean Opinion Score is a measure of voice quality using the Vocoder Mean Opinion Score Conformance Test standards document (TIA/EIA/IS 102) which is the standard subjective voice quality measurement metric. The higher score numbers represent better voice quality.

Table A

Mean Opinion Score		
Vocoder Type	Single Vocoding	Tandem Vocoding
EVRC	3.80	3.34
QCELP13	3.88	3.60
ACELP	3.91	3.67
VSELP	3.85	3.13

The vocoders listed above are more fully identified as follows:

EVRC refers to Enhanced Variable Rate Codec
QCELP refers to Qualcomm Code Excited Linear Prediction Codec
ACELP refers to Algeraic Code excited Linear Predictor
VSELP refers to Vector Sum Excited Linear Predictor.

[0011] A review of the data in Table A will clearly show that in each instance tandem vocoding of a voice signal introduces increased distortion over that of single vocoding and that this is so regardless of the type of voice coder used for the processing of voice signals.

[0012] The subjective voice quality of a given voice signal is a nonlinear function of many variables such as:

1) Voice intelligibility;
2) Voice clarity;

BNSDOCID: <EP_____1126439A2_I_>

3) Amount of background noise in the voice;
4) Digital distortion in the voice;
5) Additive noise; and
6) Frequency content of the voice (spectrum).

[0013] This invention is directed towards improving the quality of the voice generated by a voice signal subjected to tandem vocoding processing to be substantially equal to the quality of the voice generated by a voice signal that is subjected to single vocoding processing. The improvement is obtained by compensating for the digital distortion that is introduced by the second vocoding processing by modifying the frequency content of the voice signal before it is processed by the second vocoder.

[0014] Referring to Fig. 1, there is illustrated a block diagram of a mobile to desk phone call having single vocoding processing. The voice of a caller at a mobile station is processed by a voice coder (encoder) 10 and then fed to a wireless transmitter 12 for transmission to wireless receiver 14. The signal from receiver 14 is then fed to decoder 16. The output of decoder is transmitted over a path 17 that may include land lines to a desk top phone 18. The voice signal from the mobile station to the desk top phone is processed by one encoder 10 and one decoder 16 and, therefore, it is subjected to single vocoding processing.

[0015] Referring to Fig. 2, there is illustrated the block diagram of double vocoding processing. The voice of a caller at a first mobile station is processed by encoder 30,the output of which is fed to wireless transmitter 32 which transmits the signal to wireless receiver 34. The output of wireless receiver 34 is fed to decoder 36 which converts the received signal to a voice signal for transmission over the carrier's network which may include land lines 38. In this example, the originating call from the first mobile station is being sent to a second mobile station. Therefor, the voice signal from decoder 36 is sent via the carrier's network to a second encoder 40 where it is processed a second time, and then forward to wireless transmitter 42 for transmission to wireless receiver 44 located in the second or receiving mobile station. The signal from receiver 44 is fed to and processed by decoder 46, the output of which is the voice signal that is used to generate the voice heard by the person receiving the call. In Fig. 2, the call from the first mobile station to the second mobile station is encoded twice and decoded twice before it is received by the person called. Thus, in this example, the voice signal is subjected to tandem vocoding processing.

[0016] Referring to Fig. 3, there is illustrated two graphs of the frequency content of phonetically balanced voice subjected to single and tandem processing using EVRC vocoders. Phonetically balanced voice consists of a group of sentences having almost all of the phonetic transitions of the English language. The voice signal consists of 5 male and 5 female voices. Graph A is of the voice produced by single vocoding processing and Graph B is of the voice produced by tandem vocoding processing. As illustrated in Fig. 3, the voice produced by a voice signal that is subjected to tandem vocoding processing has less energy in the high frequency end compared to the voice signal produced by single vocoding processing.

[0017] This invention is directed toward improving the quality of the voice generated by a voice signal subjected to tandem vocoding by changing the spectrum of the voice signal prior to it being subjected to the second vocoding operation to compensate for digital distortion which will be generated by the second vocoding operation. This can be achieved with an adaptive filter located upstream of the second vocoder of the tandem vocoder arrangement. The adaptive filter can be a stand alone item, a part of the back end of the first decoder or a part of the front end of the second encoder.

[0018] Referring to Figs. 4, 5 and 6, there is illustrated three possible locations of the adaptive filter. Fig. 4 is a block diagram of structure in accordance with the principles of the invention where the adaptive filter is located between the first decoder and the second encoder. Fig. 5 shows the adaptive filter as a part of the back end of the decoder of the party making the call. Fig. 6 shows the adaptive filter as a part of the front end of the encoder of the party receiving the call. In each instance, the adaptive filter is located between the first decoder and the second encoder.

[0019] Fig. 7 illustrates a graph of the difference in the spectrum of one signal that is subjected to tandem vocoder processing and another signal that is subjected to single vocoding processing. A study of Fig. 7 will show that there is a marked difference between the two as the frequency increases above 3000 Hz; and, there is a smaller but possibly significant difference from 1500 Hz to 3000 Hz. An adaptive filter designed to reduce the difference above 3000 Hz will provide significant results, and further improvement can be obtained by reducing the difference from 1500 Hz to 3000 Hz. Still further improvement can be obtained by reducing the difference that occurs between 500 Hz and 1500 Hz.

[0020] Referring to Fig. 8, there is illustrated the graph of the frequency response of an adaptive filter which can be used to modify the spectrum of a voice signal fed to the second encoder.

[0021] Fig. 9 is a graph of the spectrum of a voice signal that was subjected to tandem vocoding with an adaptive filter (plot A) relative to the spectrum of a voice signal subjected to single vocoding (plot B) for one type of vocoder. Improvement of the spectrum of a signal subjected to double vocoding can be obtained by using an adaptive filter which modifies the spectrum of the voice signal from the first decoder to compensate for digital distortion which will be generated by the second decoder. Depending upon the vocoder being used, the adaptive filter can cause the spectrum

BNSDOCID: <EP_____1126439A2_I_>

of the modified signal to be close to that of the spectrum of the signal from the first vocoder for frequencies below 2400 Hz and boost the frequencies above 2400 Hz of the signal from the first vocoder by an amount from 0 to 10 db.

[0022] Fig. 10 illustrates the graph of the frequency response of an adaptive filter designed for use with two EVRC vocoders in tandem. In this embodiment, the filter is designed to pass, without change, those frequencies from the first decoder which are below 1500 Hz; to increase the db of frequencies which fall between 1500 Hz and 2400 Hz by about 2 db; to increase the db of frequencies which fall between about 2400 Hz and 2850 Hz by about 6 db; to increase the db of frequencies which fall between about 2850 Hz and 3150 Hz by about 9 db; and to increase the db of frequencies which fall between about 3150 Hz and 4000 Hz by about 6 db.

[0023] The quality of the voice generated when a voice signal was processed with single vocoding, tandem vocoding and tandem vocoding with an adaptive filter having a frequency response similar to that of Fig. 10 were evaluated using golden ear experts in the voice quality group. The results obtained indicated that the quality of the voice generated by the voice signal processed with tandem vocoding and the adaptive filter was not only better than the voice generated without the adaptive filter, but it was close to the quality of the voice generated by voice signals processed with single vocoding. The adaptive filter can be designed to modify the spectrum of the voice signal from the first vocoder to compensate for substantially all digital distortion due to the second vocoding processing.

[0024] Although the present invention has been described with respect to a typical and a preferred embodiment, it should be understood that various changes, substitutions and modifications may be suggested to one skilled in the art and it is intended that the present invention encompass such changes, substitutions and modifications as fall within the scope of the appended claims.

Claims

1. A method for communicating characterized by the steps of:

   receiving a first encoded voice signal as a first set of voice signal parameters;
   directing the first set of voice signal parameters to a first speech decoder to generate a voice signal;
   feeding the voice signal from the first speech decoder to an adaptive filter to produce a modified voice signal;
   feeding the modified voice signal to a speech encoder to convert the modified voice signal into an encoded modified voice signal represented by a second set of voice signal parameters; and
   transmitting the second set of voice signal parameters.

2. The method of claim 1 further characterized by the step of:
   modifying the spectrum of the voice signal from the first speech decoder using the adaptive filter which compensates for digital distortion which will occur when the encoded modified voice signal represented by the second set of voice signal parameters is decoded.

3. The method of claim 1 characterized in that:
   the db of frequencies above 2400 Hz of the modified voice signal have increased magnitude of from 0 to 10 db.

4. The method of claim 1 characterized in that:
   the db of frequencies below 2000 Hz of the modified voice signal have increased magnitude of from 0 to 6 db.

5. The method of claim 1 characterized in that:
   the db of frequencies above 2000 Hz of the modified voice signal have increased magnitude of from 2 db to 8 db.

6. The method of claim 1 characterized in that:
   the db of frequencies between 1500 Hz and 2400 Hz of the modified voice signal have increased magnitude of from 0 db to 2 db.

7. The method of claim 1 characterized in that:
   the db of frequencies between 2400 Hz and 2850 Hz, and those that are between 3150 H and 4000 Hz of the modified voice signal have increased magnitude of from 4 db to 8 db.

8. The method of claim 1 characterized in that:
   the db of frequencies between 2850 Hz and 3150 Hz of the modified voice signal have increased magnitude of from 7 db to 11 db.

BNSDOCID: <EP_____1126439A2_I_>

9. A wireless call connection characterized by:

a speech decoder coupled to a wireless receiver for receiving a signal made up of a set of parameters representative of a voice signal and for converting the received signal to a voice signal;
an adaptive filter coupled to receive the voice signal from the speech decoder to produce a modified voice signal which, when encoded, will produce a second signal made up of another set of parameters representative of the voice signal.

10. The structure of claim 9 characterized in that the adaptive filter modifies the magnitude of selective frequencies of the spectrum of the voice signal from the decoder.

11. The structure of claim 10 characterized in that the adaptive filter modifies the magnitude of selective frequencies of the spectrum of the voice signal from the decoder to compensate for digital distortion caused by encoding and decoding the modified voice signal.

12. The structure of claim 10 characterized in that the adaptive filter increases the db of frequencies above 3000 Hz of the voice signal from the decoder to compensate for digital distortion caused by encoding and decoding the modified voice signal from the adaptive filter.

13. The structure of claim 10 characterized in that the adaptive filter increases the db of frequencies above 1500 Hz of the voice signal from the decoder to compensate for digital distortion caused by encoding and decoding the modified voice signal from the adaptive filter.

14. The structure of claim 10 characterized in that the adaptive filter increases the db of frequencies above 300 Hz of the voice signal from the decoder to compensate for digital distortion caused by encoding and decoding the modified voice signal from the adaptive filter.

15. The structure of claim 9 characterized in that the adaptive filter increases the db of selective frequencies of the spectrum of the voice signal from the decoder to cause the spectrum of a voice signal generated by subsequent encoding and decoding of the modified signal to be close to that of the voice signal from the speech decoder.

BNSDOCID: <EP_____1126439A2_I_>

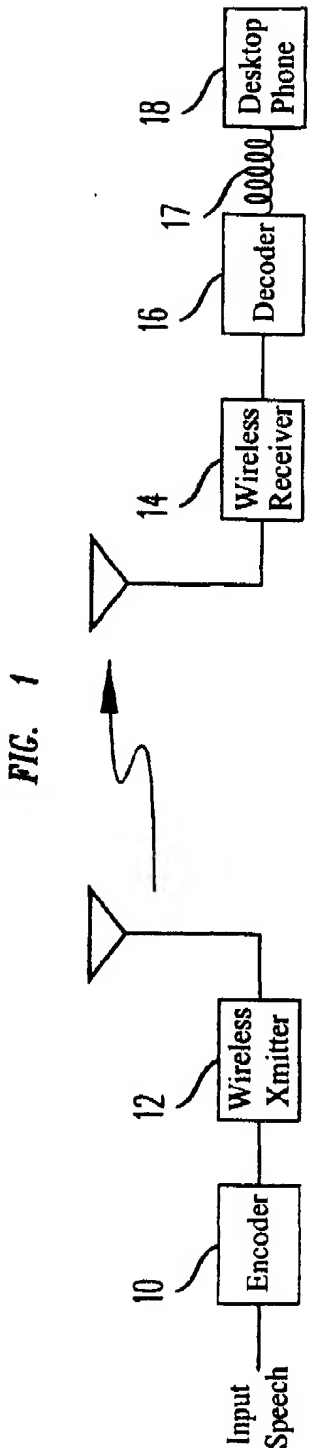


FIG. 1

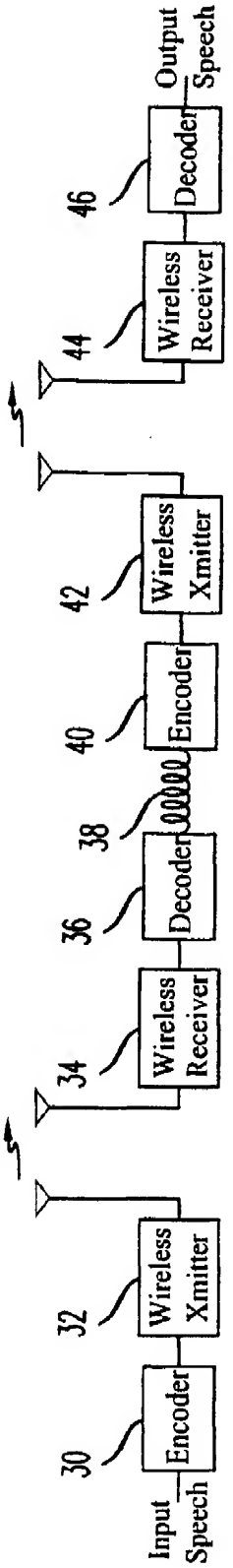
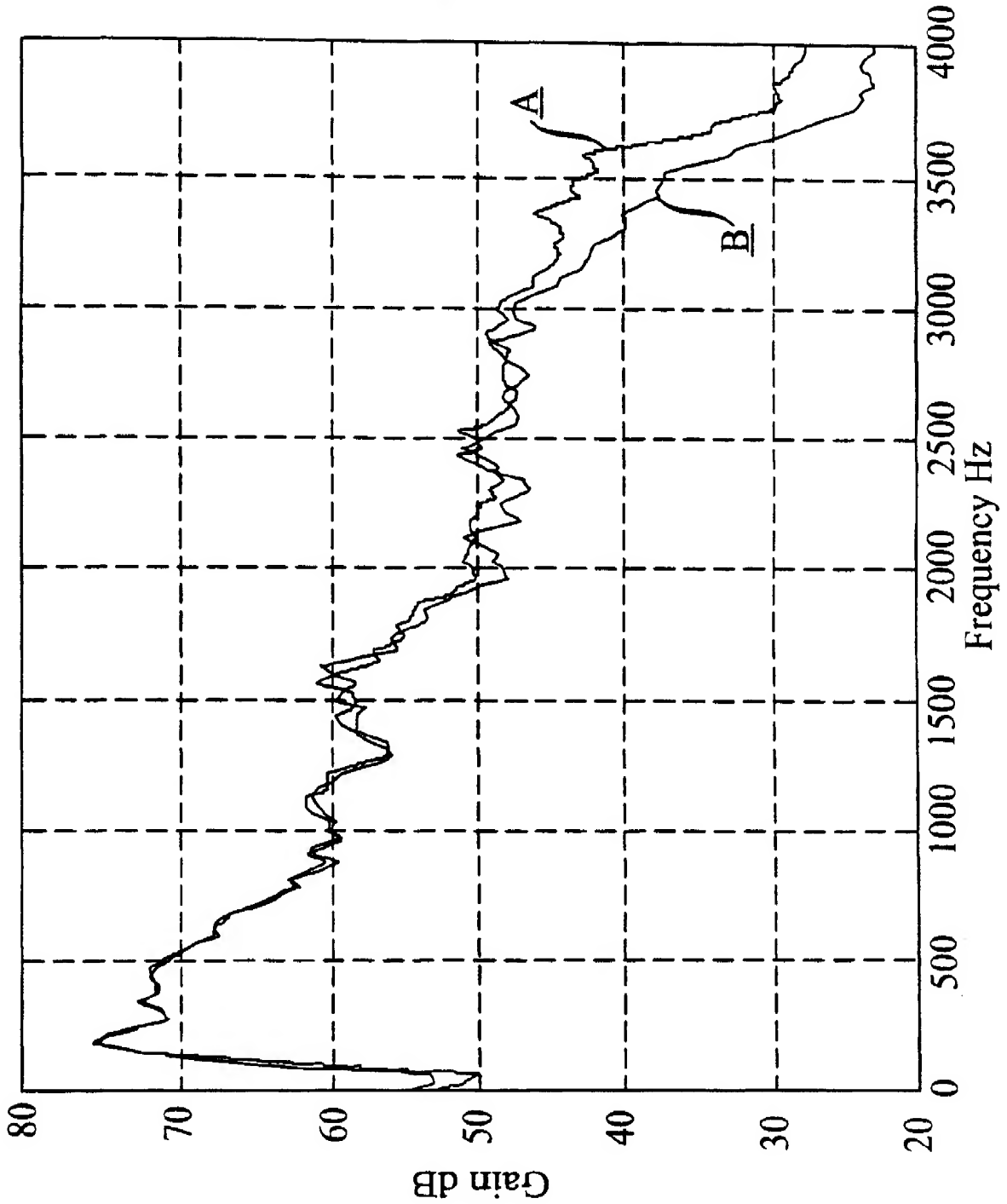


FIG. 2

BNSDOCID: <EP_____1126439A2_I_>

FIG. 3



BNSDOCID: <EP_____1126439A2_I_>

FIG. 4

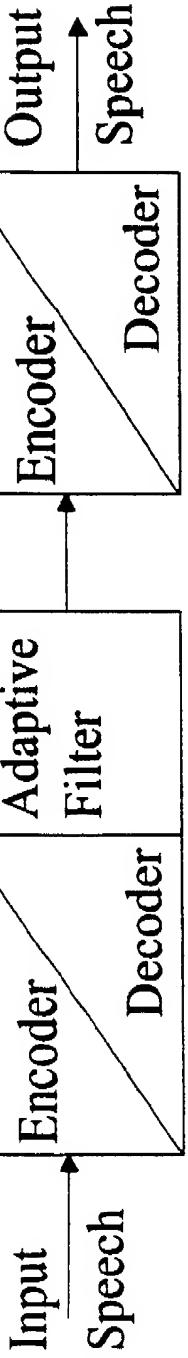


FIG. 5

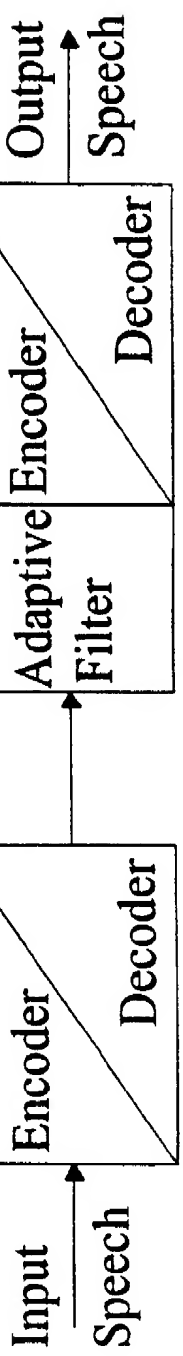
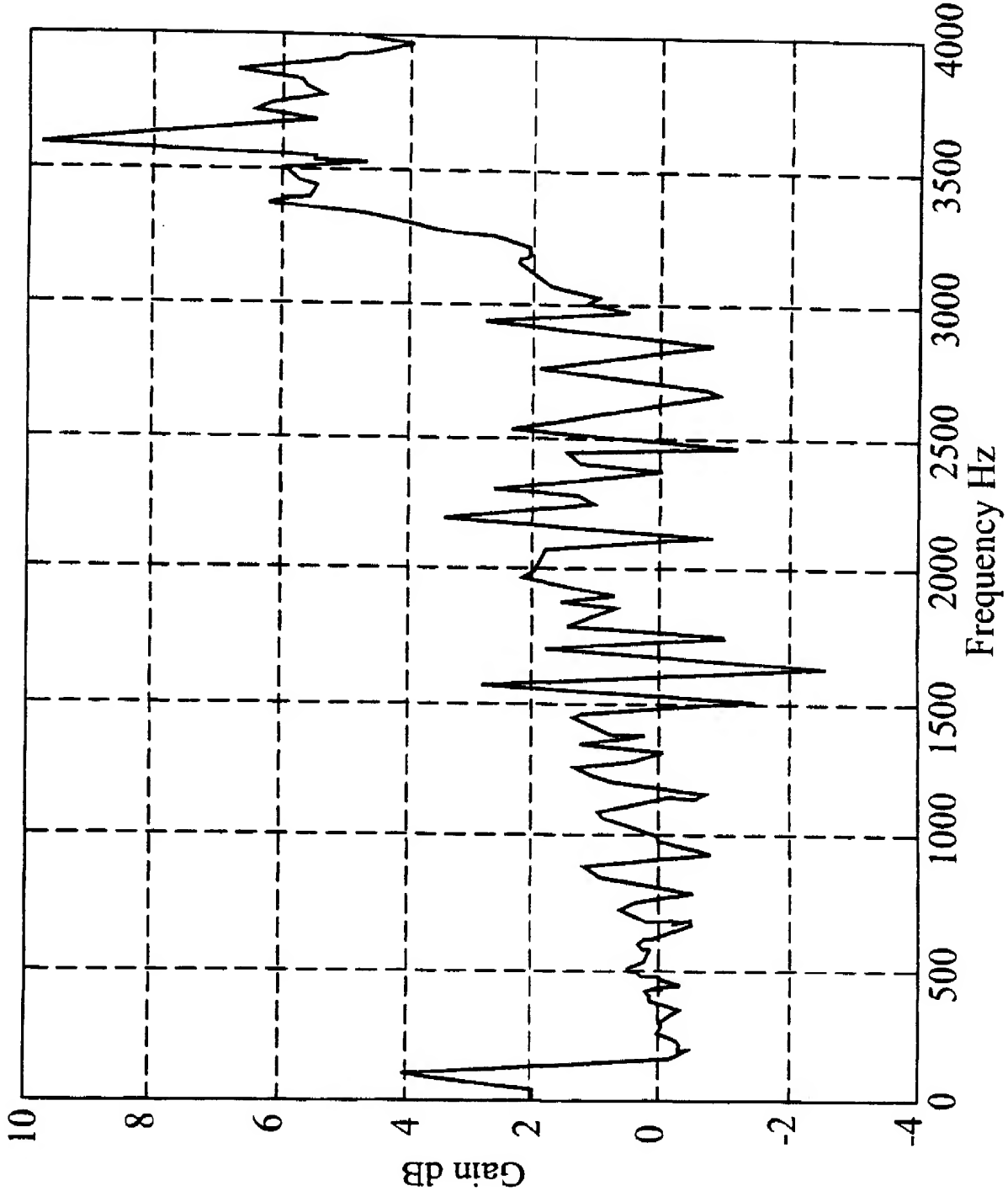


FIG. 6

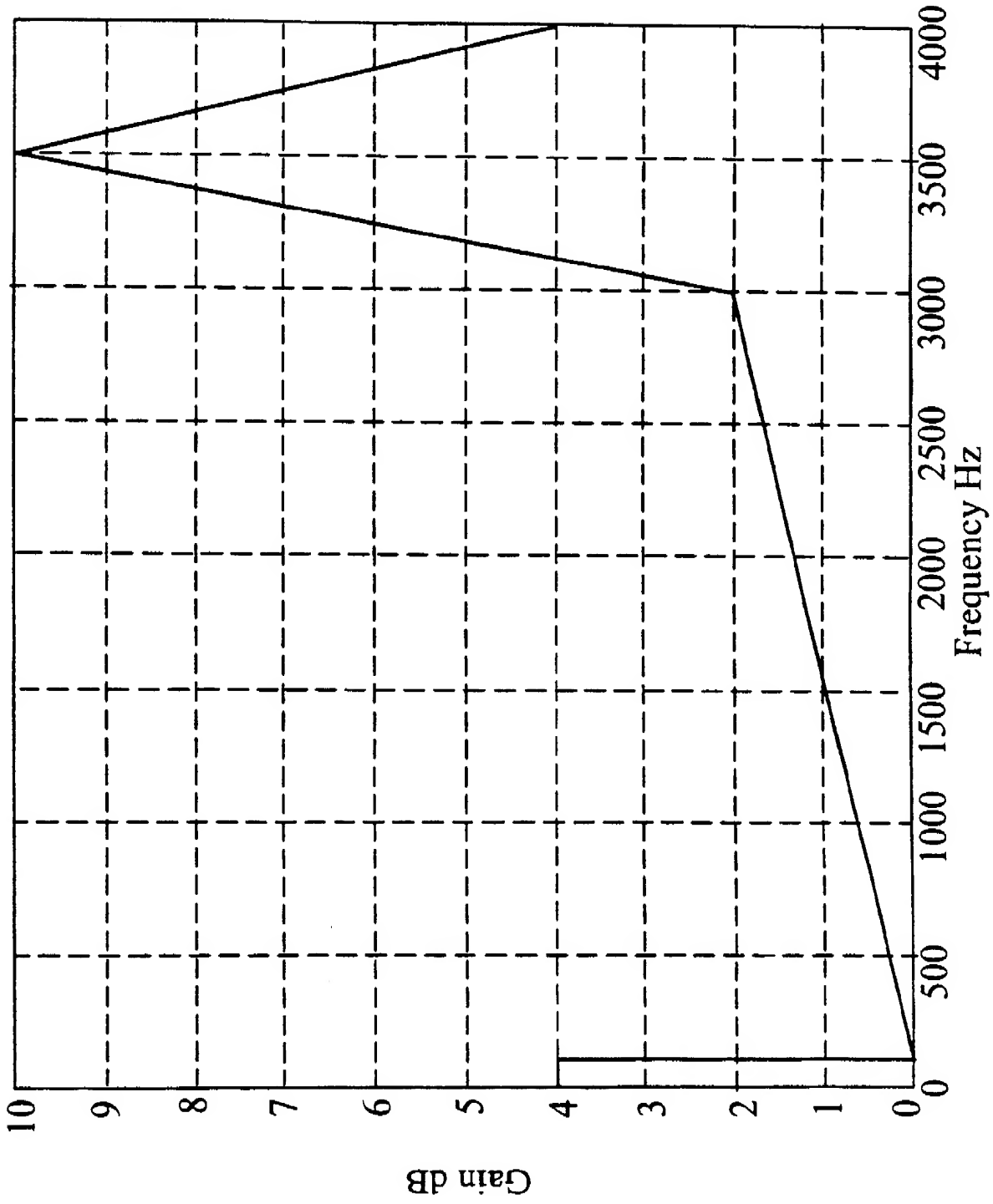
BNSDOCID: <EP_____1126439A2_I_>

FIG. 7



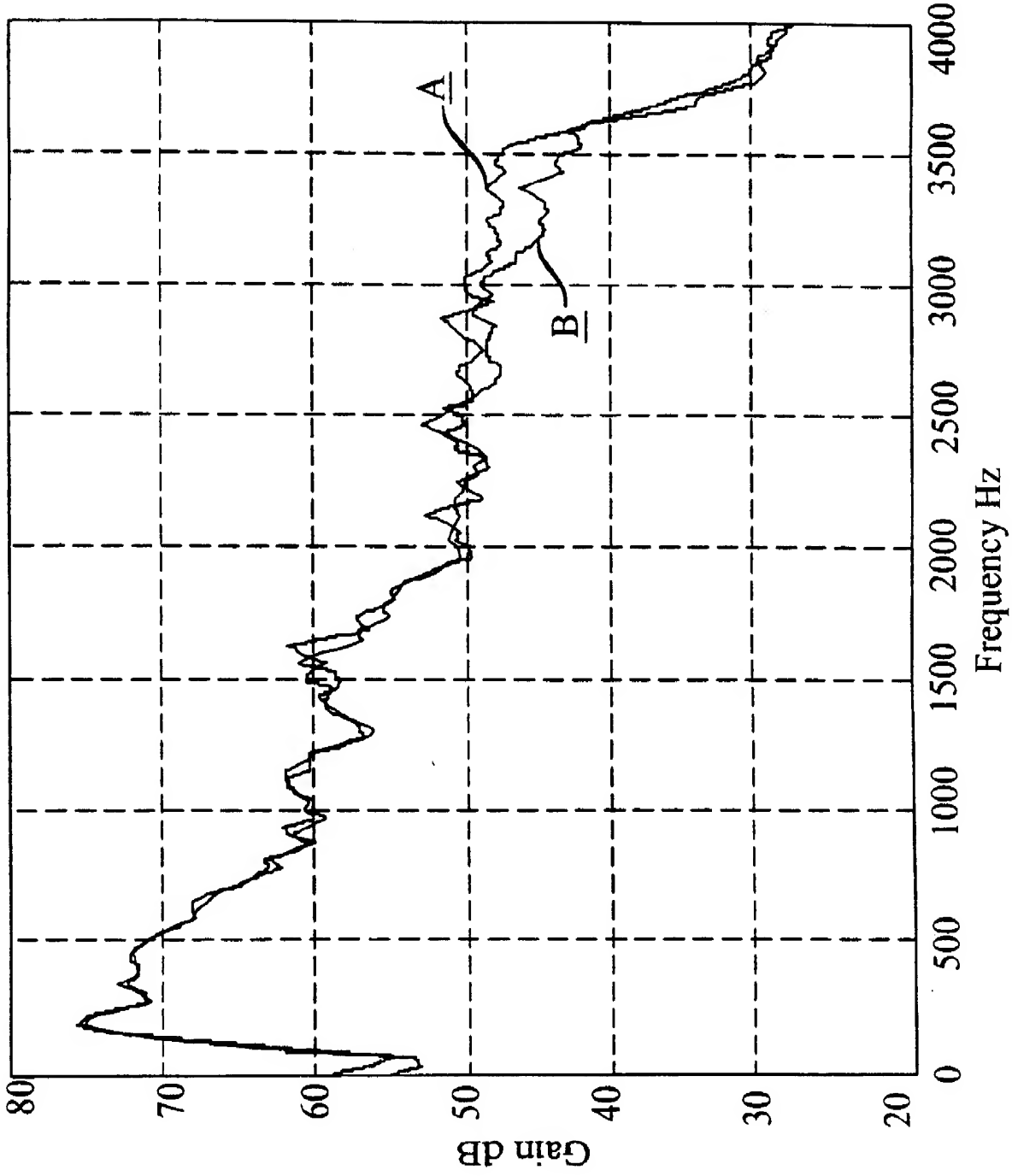
BNSDOCID: <EP_____1126439A2_I_>

FIG. 8



BNSDOCID: <EP_____1126439A2_I_>

FIG. 9



BNSDOCID: <EP_____1126439A2_I_>

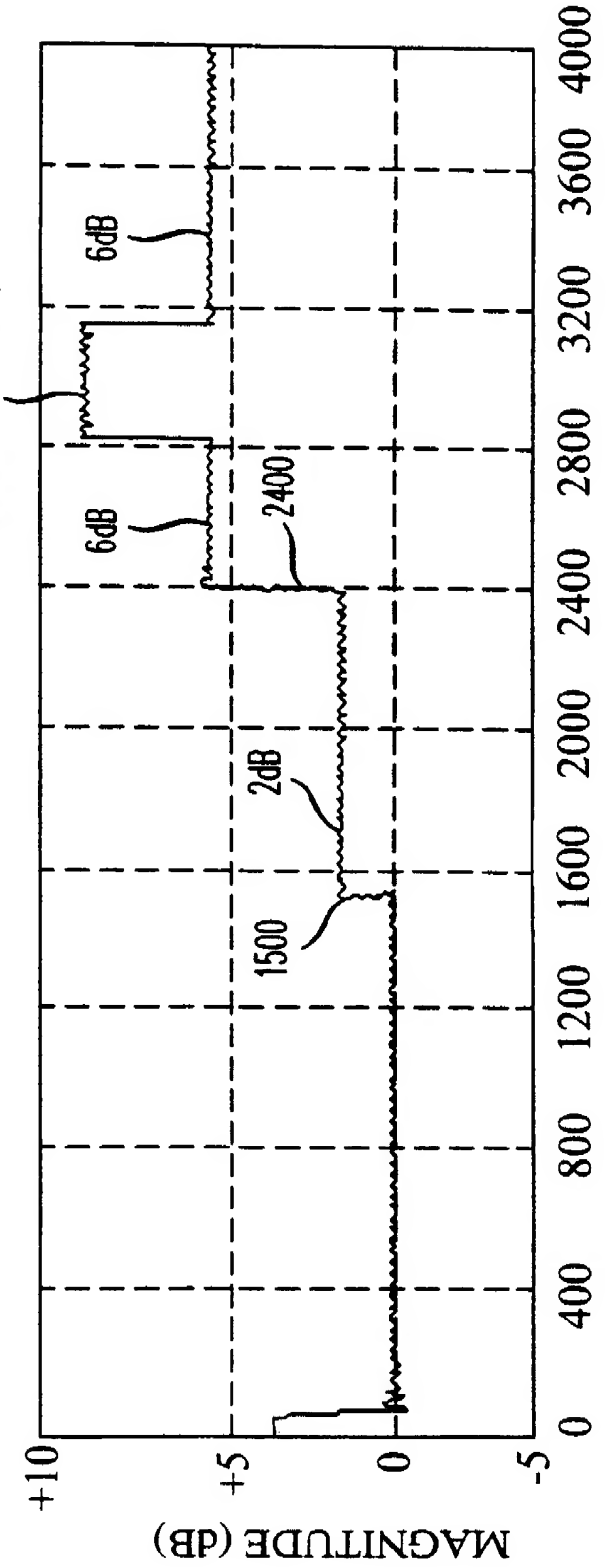


FIG. 10

BNSDOCID: <EP_____1126439A2_I_>